Bayesian Networks: Inference and Independence

Bayesian Networks

Today:

- Bayesian Networks
- How do we perform inference on Bayesian Networks?
- How do we reason about independence in Bayesian Networks?

Review

Joint

$P(A, B, C)$

$P(A=1, B=3, C=5)$

Conditional

$P(X \mid Y)$ ← distribution-valued function of Y

$P(X=1 \mid Y=1)$

Marginal

$P(X)$

Independence

$P(X, Y) = P(X)P(Y)$

$X \perp Y$

$P(X \mid Y) = P(X)$

Conditional Indep.

$P(X, Y \mid Z) = P(X \mid Z)P(Y \mid Z)$

$X \perp Y \mid Z$

$P(X \mid Z) = P(X \mid Y, Z)$

Joint Distribution Complexity

Binary Random Variables X_1, X_2, X_3

X_0 $\quad P(X_0)$

X_0	
0	0.7
1	0.3

How many independent parameters (θ) to specify joint distribution?

$2^3 = 8$

X_1	X_2	X_3	$P(X_1, X_2, X_3)$
0	0	0	
0	0	1	
0	1	0	
	⋮		

Joint Distribution Complexity

Binary Random Variables X_1, X_2, X_3

How many independent parameters (θ) to specify joint distribution? 7

Joint Distribution Complexity

Binary Random Variables X_1, X_2, X_3

How many independent parameters (θ) to specify joint distribution? 7

For n binary R.V.s, $2^n - 1$ independent parameters specify the joint distribution.

Joint Distribution Complexity

Binary Random Variables X_1, X_2, X_3

How many independent parameters (θ) to specify joint distribution? 7

For n binary R.V.s, $2^n - 1$ independent parameters specify the joint distribution.

In general

vector of parameters

$$\dim(\theta) = \prod_{i=1}^{n} |\text{support}(X_i)| - 1$$

Bayesian Network

Full Story: "Causality: Models, Reasoning and Inference" by Judea Pearl

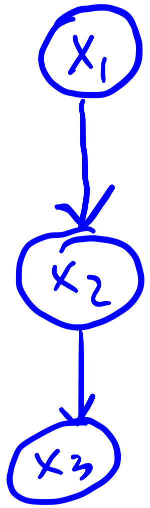
Bayesian Network

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**

Full Story: "Causality: Models, Reasoning and Inference" by Judea Pearl

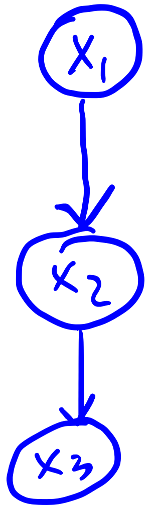
Bayesian Network

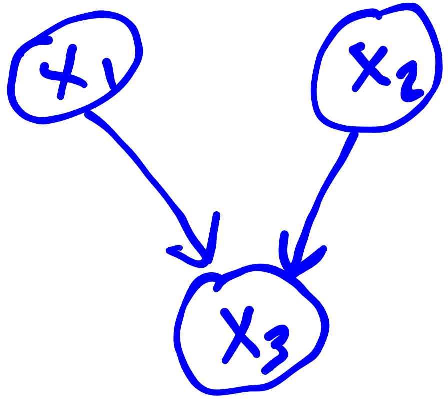
Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**



Bayesian Network

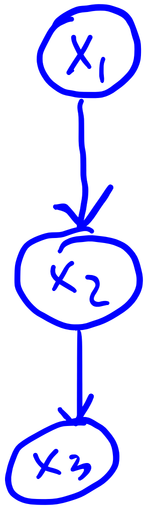
Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**

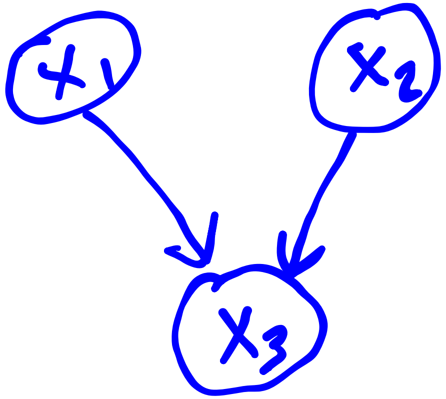




Bayesian Network

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**



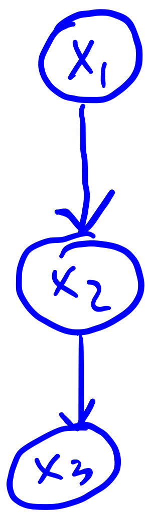


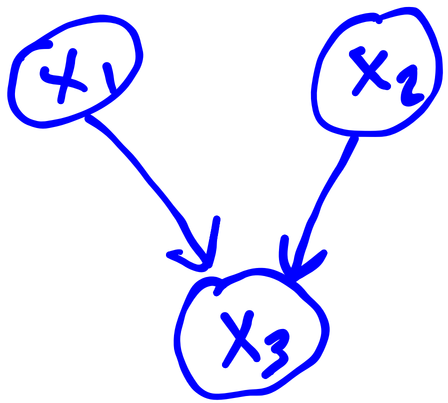
- Node:
- Edges encode:

Full Story: "Causality: Models, Reasoning and Inference" by Judea Pearl

Bayesian Network

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**





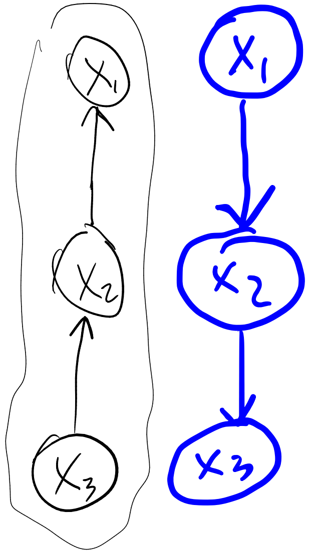
- Node: Random Variable
- Edges encode:

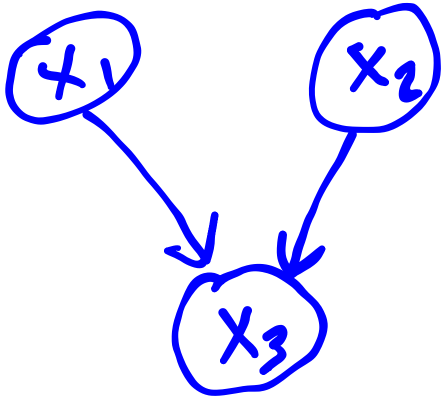
Full Story: "Causality: Models, Reasoning and Inference" by Judea Pearl

Bayesian Network

Bayesian Network: Directed Acyclic Graph (DAG) that represents a **joint probability distribution**

$$P(X_1, X_2, X_3) = P(X_1)\,P(X_2)\,P(X_3 \mid X_1, X_2)$$





- Node: Random Variable
- Edges encode:

$$P(X_{1:n}) = \prod_{i=1}^{n} P(X_i \mid \mathrm{pa}(X_i))$$

Joint

$$P(X_1, X_2, X_3) = P(X_1)\,P(X_2 \mid X_1)\,P(X_3 \mid X_2)$$

X_1, X_2, X_3	$P(X_1, X_2, X_3)$

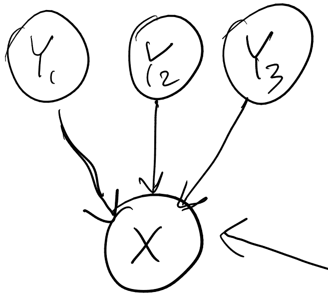
Full Story: "Causality: Models, Reasoning and Inference" by Judea Pearl

Counting Parameters

For discrete R.V.s:

$$\dim(\theta_X) = \underline{(|\text{support}(X)| - 1)} \underline{\prod_{Y \in Pa(X)} |\text{support}(Y)|}$$

Counting Parameters



For discrete R.V.s:

$$\dim(\theta_X) = \underbrace{(|\text{support}(X)| - 1)}_{1}\underbrace{\prod_{Y \in Pa(X)} \underbrace{|\text{support}(Y)|}_{2}}_{4}$$

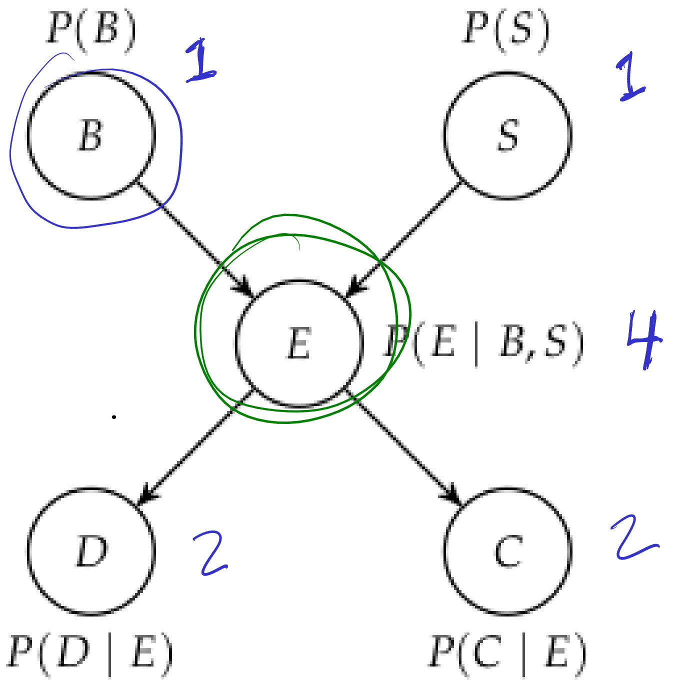
support(B) = {0,1}

|support(B)| = 2

$P(X \mid Y_1, Y_2, Y_3)$

0 0 0 → $P(X \mid Y_1 = 0, Y_2 = 0, Y_3 = 0)$ ⇒ 1 param

0 0 1 ⇒ 1 param



support(E) = {0,1}

| | = 2

Pa(E) = {B, S}

|support(B)| = 2

|support(S)| = 2

B.N.

Number of parameters: 10

Number of parameters for naive representation $2^5 - 1 = 31$

Inference

Inputs

Outputs

Inference

Inputs

- Bayesian network structure

Outputs

Inference

Inputs

- Bayesian network structure
- Bayesian network parameters

Outputs

Inference

Inputs

- Bayesian network structure
- Bayesian network parameters
- Values of *evidence variables*

Outputs

Inference

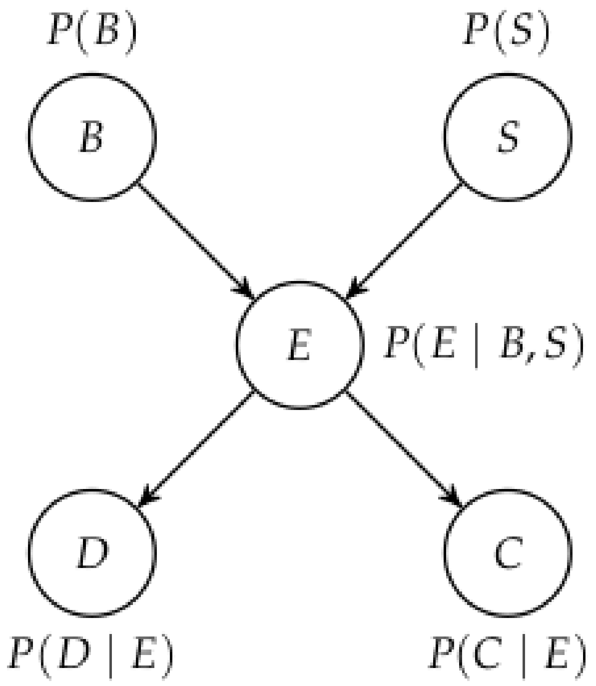
Inputs

- Bayesian network structure
- Bayesian network parameters
- Values of *evidence variables*

Outputs

- Posterior distribution of *query variables*

Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

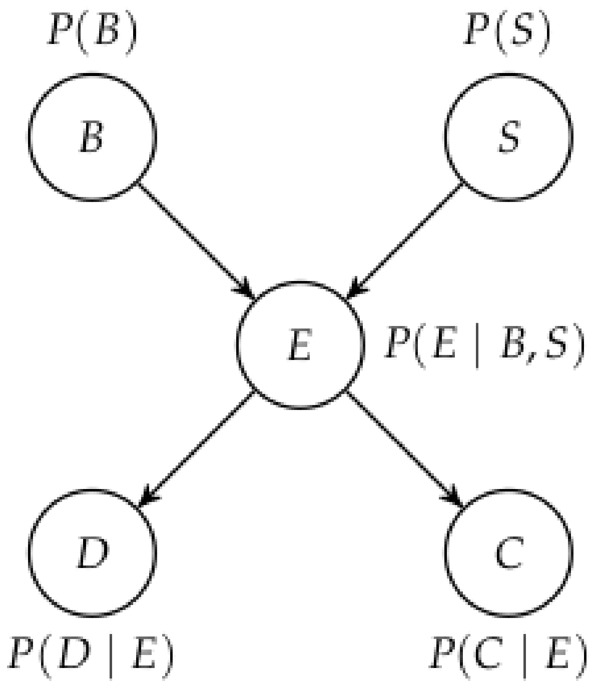
Inputs

- Bayesian network structure
- Bayesian network parameters
- Values of *evidence variables*

Outputs

- Posterior distribution of *query variables*

Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Inputs

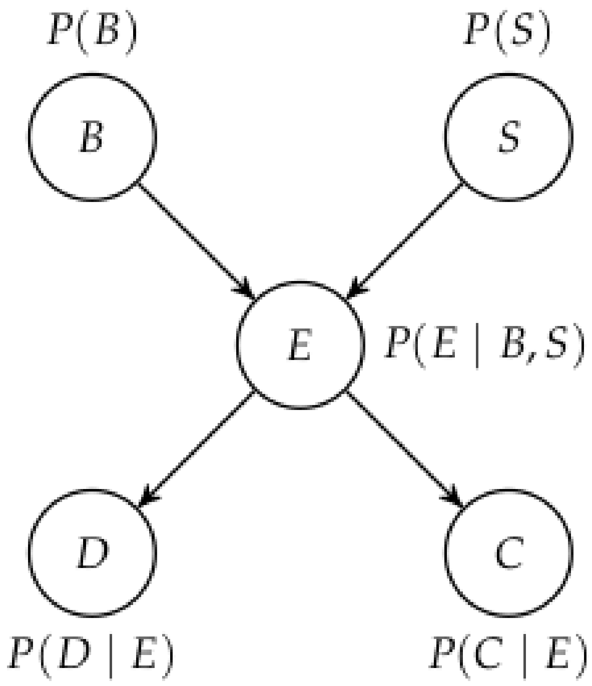
- Bayesian network structure
- Bayesian network parameters
- Values of *evidence variables*

Outputs

- Posterior distribution of *query variables*

Given that you have detected a trajectory deviation, and the battery has not failed what is the probability of a solar panel failure?

Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Inputs

- Bayesian network structure
- Bayesian network parameters
- Values of *evidence variables*

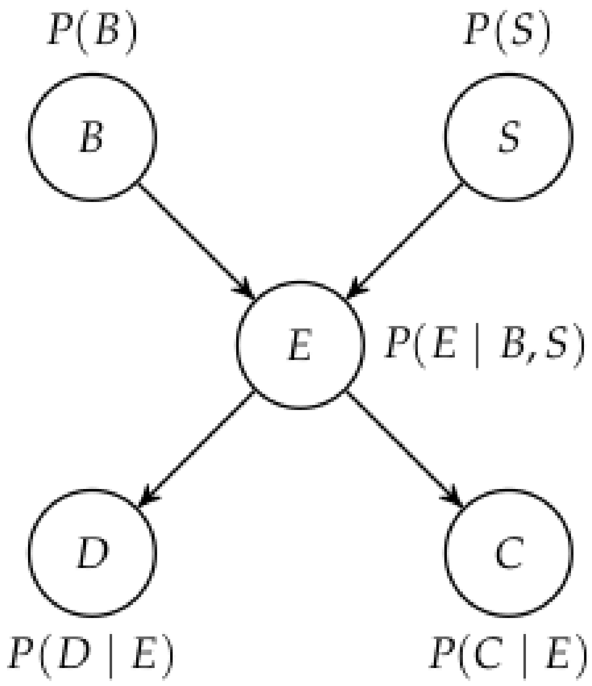
Outputs

- Posterior distribution of *query variables*

Given that you have detected a trajectory deviation, and the battery has not failed what is the probability of a solar panel failure?

$$\underbrace{P(S = 1}_{\text{query}} \mid \underbrace{D = 1, B = 0)}_{\text{Evidence}}$$

Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Inputs

- Bayesian network structure
- Bayesian network parameters
- Values of *evidence variables*

Outputs

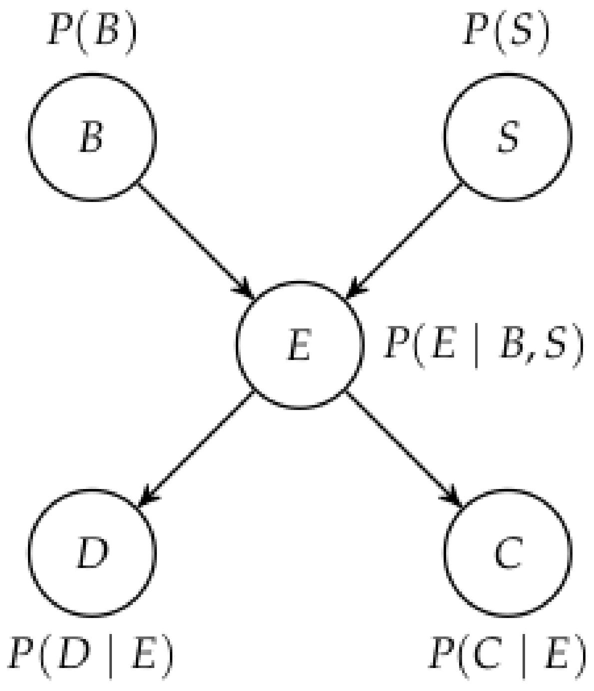
- Posterior distribution of *query variables*

Given that you have detected a trajectory deviation, and the battery has not failed what is the probability of a solar panel failure?

$$P(S = 1 \mid D = 1, B = 0)$$

Exact

Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Inputs

- Bayesian network structure
- Bayesian network parameters
- Values of *evidence variables*

Outputs

- Posterior distribution of *query variables*

Given that you have detected a trajectory deviation, and the battery has not failed what is the probability of a solar panel failure?

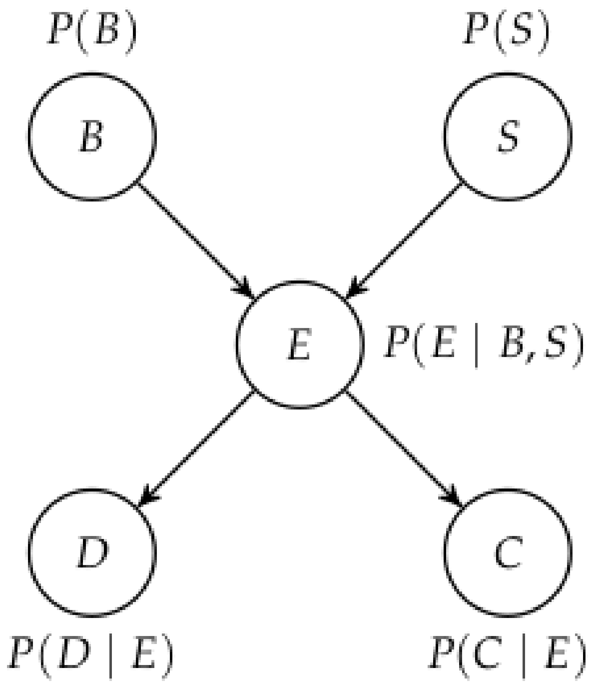
$$P(S = 1 \mid D = 1, B = 0)$$

Exact

Approximate

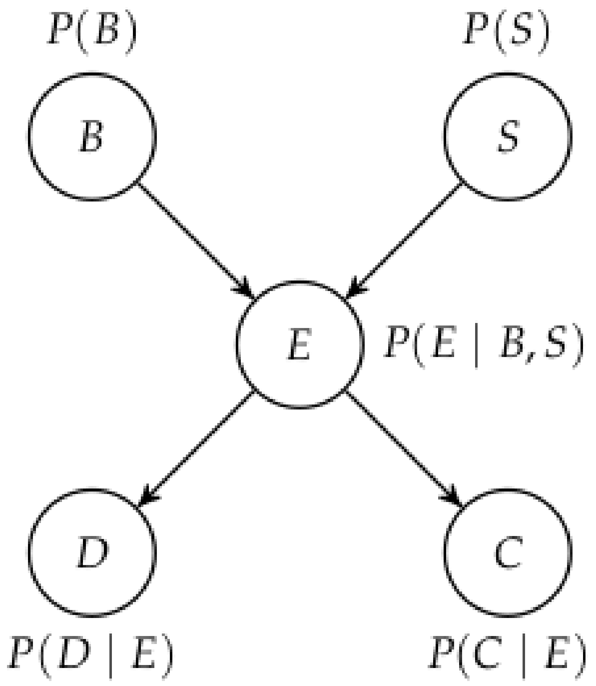
Exact Inference

Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

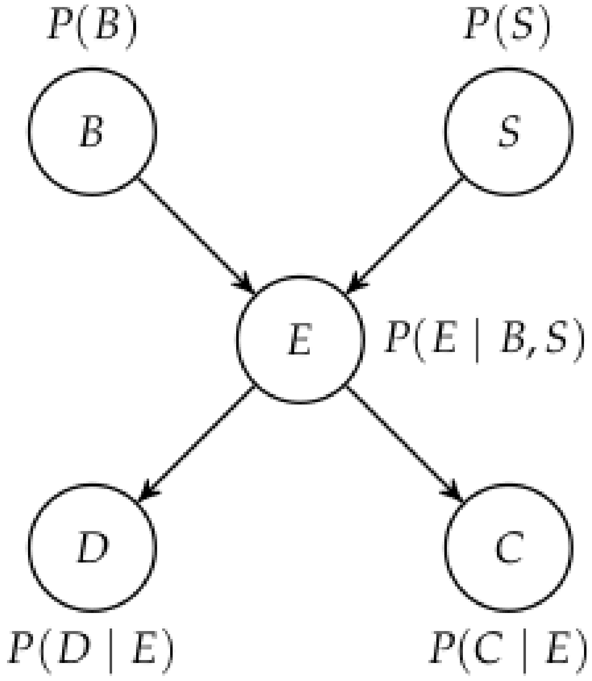
Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

$$P(S{=}1 \mid D{=}1, B{=}0)$$

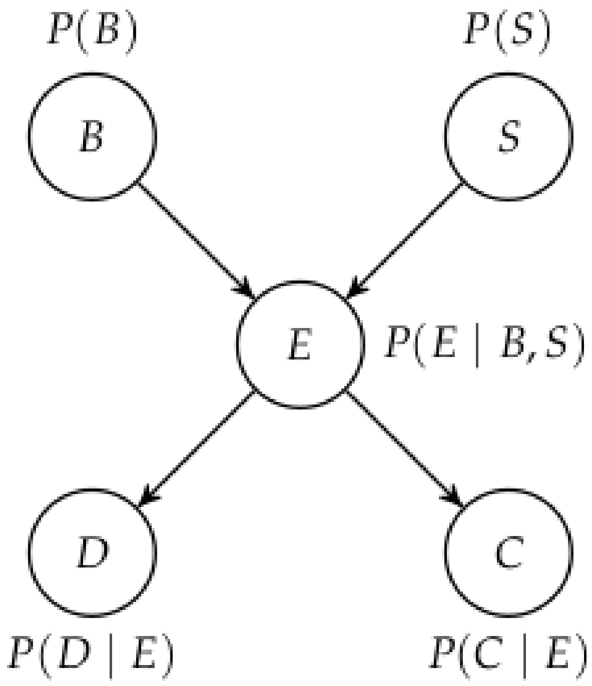
Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

$$P(S{=}1 \mid D{=}1, B{=}0) \;=\; \frac{P(S{=}1, D{=}1, B{=}0)}{P(D{=}1, B{=}0)}$$

Exact Inference

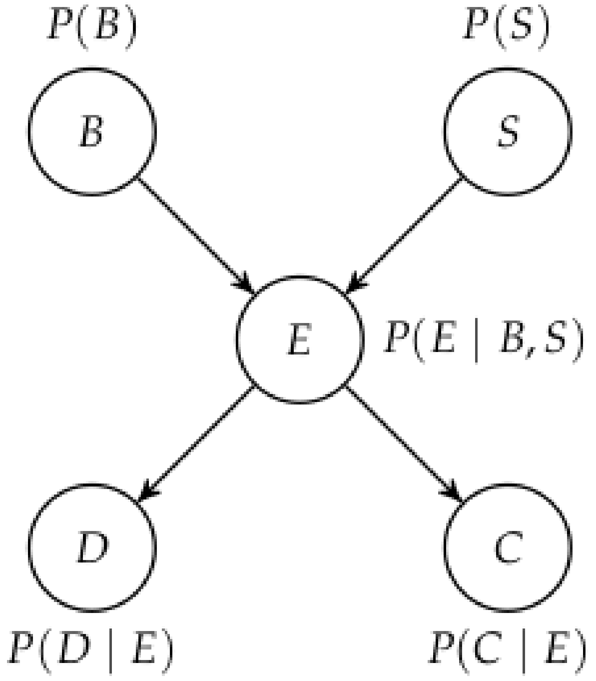


B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

$$P(S{=}1 \mid D{=}1, B{=}0) = \frac{P(S{=}1, D{=}1, B{=}0)}{P(D{=}1, B{=}0)}$$

$$P(S{=}1, D{=}1, B{=}0) = \sum_{\underline{e,c}} P(B{=}0, S{=}1, E{=}e, D{=}1, C{=}c)$$

Exact Inference



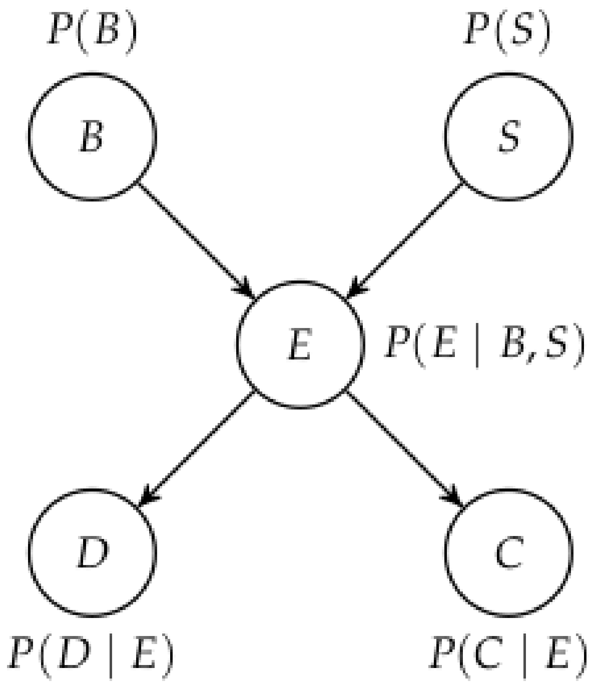
B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

$$P(S{=}1 \mid D{=}1, B{=}0) = \frac{P(S{=}1, D{=}1, B{=}0)}{P(D{=}1, B{=}0)}$$

$$P(S{=}1, D{=}1, B{=}0) = \sum_{e,c} P(B{=}0, S{=}1, E{=}e, D{=}1, C{=}c)$$

$$P(B{=}0, S{=}1, E, D{=}1, C)$$

Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

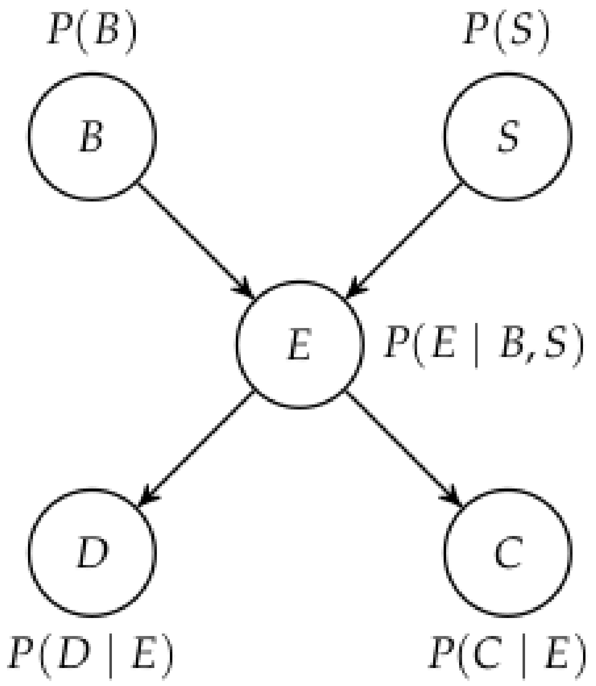
$$P(S{=}1 \mid D{=}1, B{=}0) = \frac{P(S{=}1, D{=}1, B{=}0)}{P(D{=}1, B{=}0)}$$

$$P(S{=}1, D{=}1, B{=}0) = \sum_{e,c} P(B{=}0, S{=}1, E{=}e, D{=}1, C{=}c)$$

$$P(B{=}0, S{=}1, E, D{=}1, C)$$
$$= P(B{=}0)\,P(S{=}1)\,P(E \mid B{=}0, S{=}1)\,P(D{=}1 \mid E)\,P(C{=}1 \mid E)$$

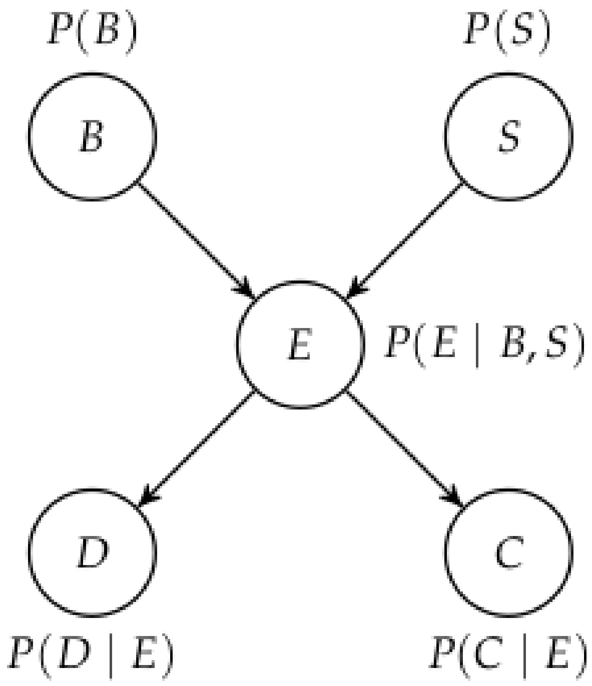
Bayesian Network

Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

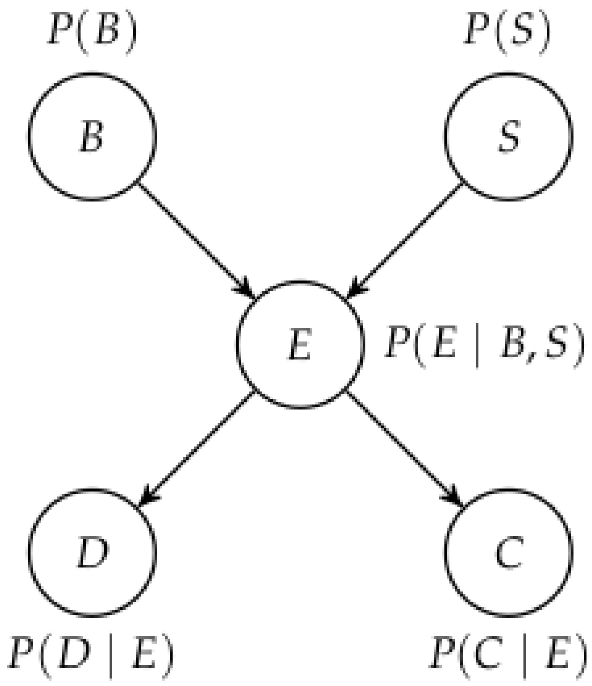
Product

X	Y	$\phi_1(X, Y)$
0	0	0.3
0	1	0.4
1	0	0.2
1	1	0.1

Y	Z	$\phi_2(Y, Z)$
0	0	0.2
0	1	0.0
1	0	0.3
1	1	0.5

X	Y	Z	$\phi_3(X, Y, Z)$
0	0	0	0.06
0	0	1	0.00
0	1	0	0.12
0	1	1	0.20
1	0	0	0.04
1	0	1	0.00
1	1	0	0.03
1	1	1	0.05

Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Product

X	Y	$\phi_1(X,Y)$
0	0	0.3
0	1	0.4
1	0	0.2
1	1	0.1

Y	Z	$\phi_2(Y,Z)$
0	0	0.2
0	1	0.0
1	0	0.3
1	1	0.5

X	Y	Z	$\phi_3(X,Y,Z)$
0	0	0	0.06
0	0	1	0.00
0	1	0	0.12
0	1	1	0.20
1	0	0	0.04
1	0	1	0.00
1	1	0	0.03
1	1	1	0.05

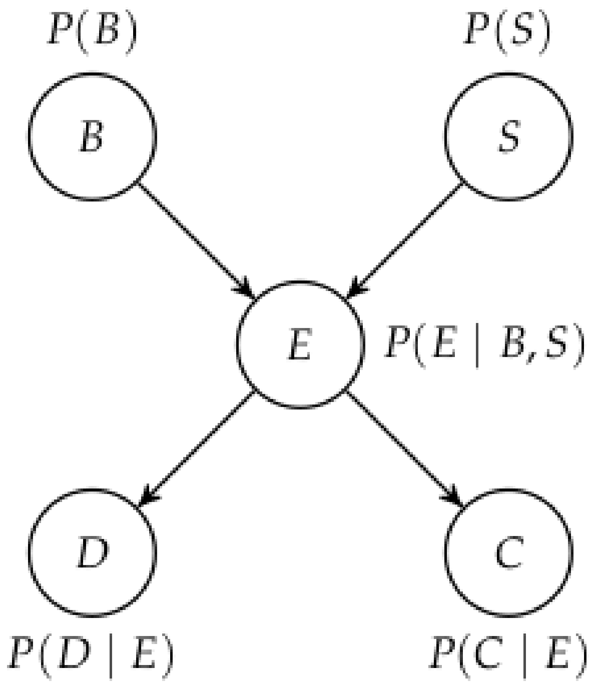
Condition

X	Y	Z	$\phi(X,Y,Z)$
0	0	0	0.08
0	0	1	0.31
0	1	0	0.09
0	1	1	0.37
1	0	0	0.01
1	0	1	0.05
1	1	0	0.02
1	1	1	0.07

$Y = 1$

X	Z	$\phi(X,Z)$
0	0	0.09
0	1	0.37
1	0	0.02
1	1	0.07

Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Product

X	Y	$\phi_1(X,Y)$
0	0	0.3
0	1	0.4
1	0	0.2
1	1	0.1

Y	Z	$\phi_2(Y,Z)$
0	0	0.2
0	1	0.0
1	0	0.3
1	1	0.5

X	Y	Z	$\phi_3(X,Y,Z)$
0	0	0	0.06
0	0	1	0.00
0	1	0	0.12
0	1	1	0.20
1	0	0	0.04
1	0	1	0.00
1	1	0	0.03
1	1	1	0.05

Condition

X	Y	Z	$\phi(X,Y,Z)$
0	0	0	0.08
0	0	1	0.31
0	1	0	0.09
0	1	1	0.37
1	0	0	0.01
1	0	1	0.05
1	1	0	0.02
1	1	1	0.07

$Y = 1$

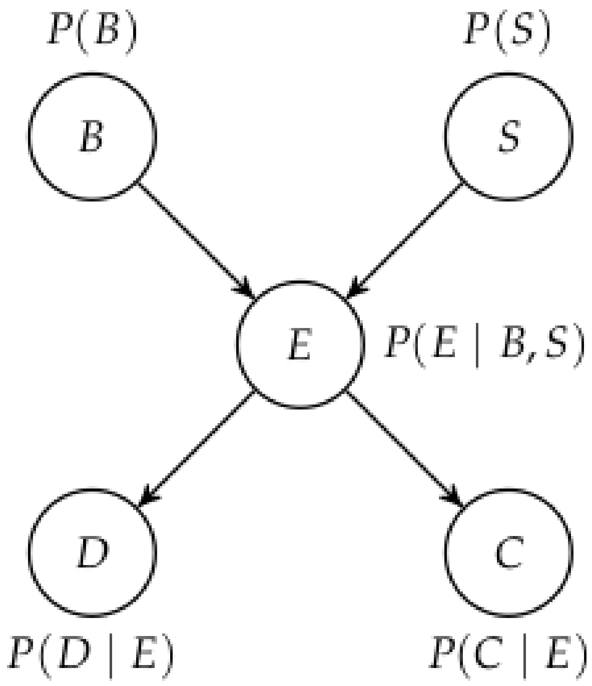
X	Z	$\phi(X,Z)$
0	0	0.09
0	1	0.37
1	0	0.02
1	1	0.07

Marginalize

X	Y	Z	$\phi(X,Y,Z)$
0	0	0	0.08
0	0	1	0.31
0	1	0	0.09
0	1	1	0.37
1	0	0	0.01
1	0	1	0.05
1	1	0	0.02
1	1	1	0.07

X	Z	$\phi(X,Z)$
0	0	0.17
0	1	0.68
1	0	0.03
1	1	0.12

Exact Inference



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

Product

X	Y	$\phi_1(X,Y)$
0	0	0.3
0	1	0.4
1	0	0.2
1	1	0.1

Y	Z	$\phi_2(Y,Z)$
0	0	0.2
0	1	0.0
1	0	0.3
1	1	0.5

X	Y	Z	$\phi_3(X,Y,Z)$
0	0	0	0.06
0	0	1	0.00
0	1	0	0.12
0	1	1	0.20
1	0	0	0.04
1	0	1	0.00
1	1	0	0.03
1	1	1	0.05

Condition

X	Y	Z	$\phi(X,Y,Z)$
0	0	0	0.08
0	0	1	0.31
0	1	0	0.09
0	1	1	0.37
1	0	0	0.01
1	0	1	0.05
1	1	0	0.02
1	1	1	0.07

$Y = 1$

X	Z	$\phi(X,Z)$
0	0	0.09
0	1	0.37
1	0	0.02
1	1	0.07

Marginalize

X	Y	Z	$\phi(X,Y,Z)$
0	0	0	0.08
0	0	1	0.31
0	1	0	0.09
0	1	1	0.37
1	0	0	0.01
1	0	1	0.05
1	1	0	0.02
1	1	1	0.07

X	Z	$\phi(X,Z)$
0	0	0.17
0	1	0.68
1	0	0.03
1	1	0.12

```
struct ExactInference end

function infer(M::ExactInference, bn, query, evidence)
    ϕ = prod(bn.factors)
    ϕ = condition(ϕ, evidence)
    for name in setdiff(variablenames(ϕ), query)
        ϕ = marginalize(ϕ, name)
    end
    return normalize!(ϕ)
end
```

Exact Inference

```julia
struct ExactInference end

function infer(M::ExactInference, bn, query, evidence)
    ϕ = prod(bn.factors)
    ϕ = condition(ϕ, evidence)
    for name in setdiff(variablenames(ϕ), query)
        ϕ = marginalize(ϕ, name)
    end
    return normalize!(ϕ)
end
```

```julia
struct VariableElimination
    ordering # array of variable indices
end

function infer(M::VariableElimination, bn, query, evidence)
    Φ = [condition(ϕ, evidence) for ϕ in bn.factors]
    for i in M.ordering
        name = bn.vars[i].name
        if name ∉ query
            inds = findall(ϕ->in_scope(name, ϕ), Φ)
            if !isempty(inds)
                ϕ = prod(Φ[inds])
                deleteat!(Φ, inds)
                ϕ = marginalize(ϕ, name)
                push!(Φ, ϕ)
            end
        end
    end
    return normalize!(prod(Φ))
end
```

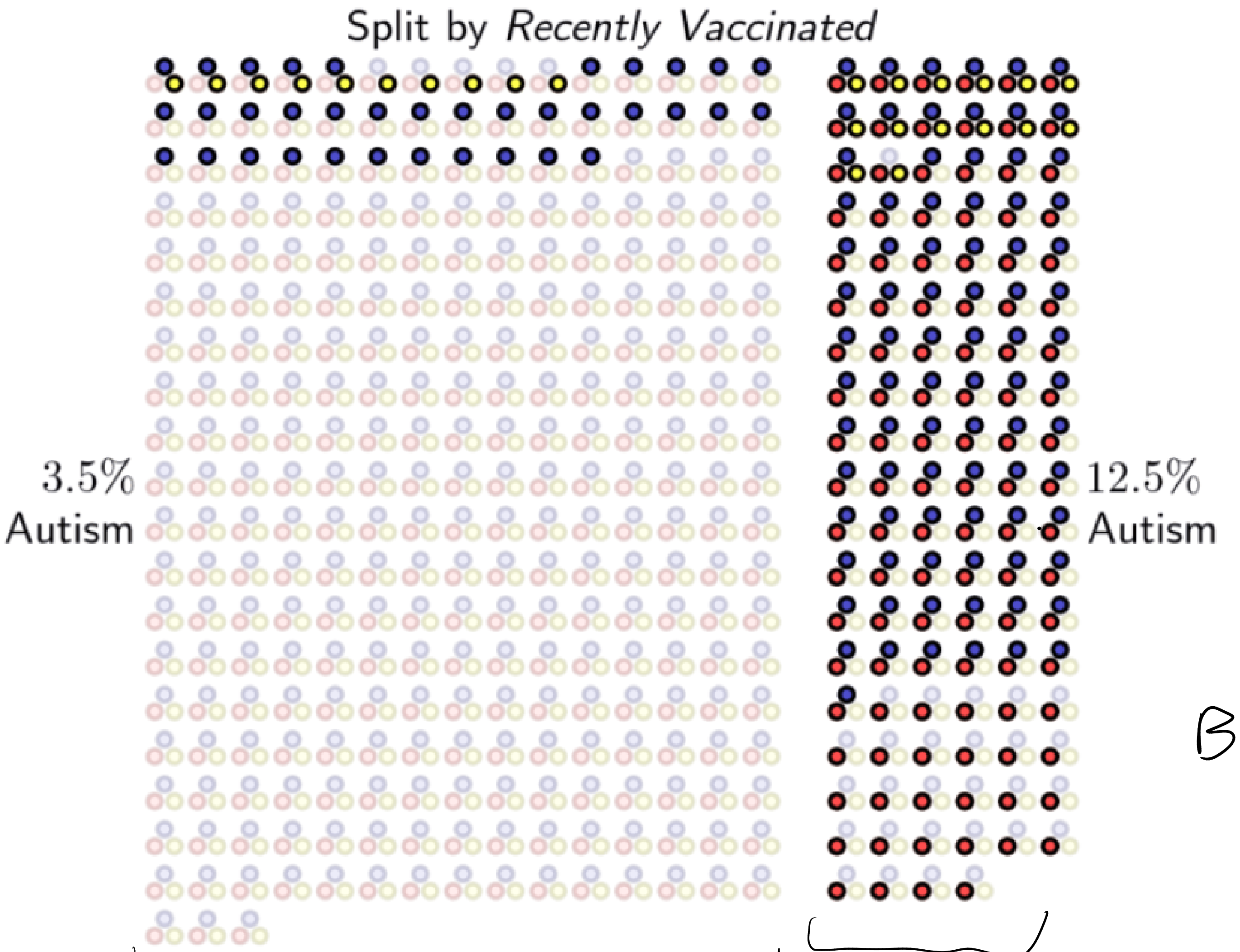
product over smaller number of variables

choosing order to eliminate variables is difficult (NP-hard)

Break

Yellow: Autism

Red: recently vaccinated



Does this imply a link between autism and vaccination

$P(A)$ $P(V)$

$P(A,V) \neq P(A)P(V)$

$A \not\perp V$

Blue: Child

$$\begin{tikzcd} & C \arrow[dl] \arrow[dr] & \\ A & & V \end{tikzcd}$$

$A \perp V \mid C$

What does conditional independence mean?

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z$

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$$X \perp Y \mid Z \implies$$

https://kunalmenda.com/2019/02/21/causation-and-correlation/

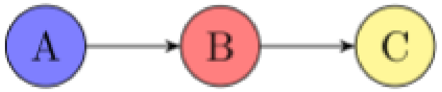
What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

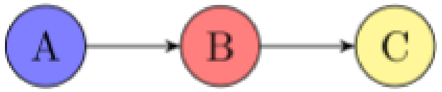
$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



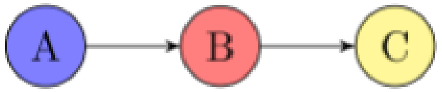
$A \perp C \mid B$?

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z

$P(X \mid Z) = P(X \mid Y, Z)$

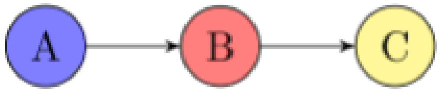


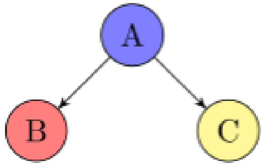
$A \perp C \mid B$? Yes

https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



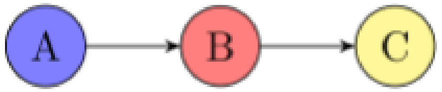


$A \perp C \mid B$? Yes

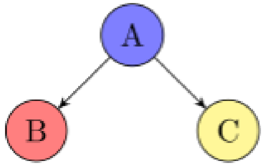
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



$A \perp C \mid B$? Yes

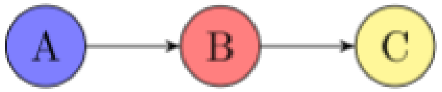


$B \perp C \mid A$?

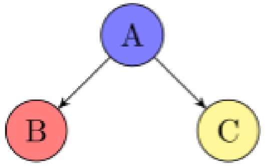
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



$A \perp C \mid B$? Yes

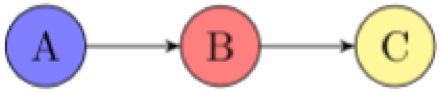


$B \perp C \mid A$? Yes

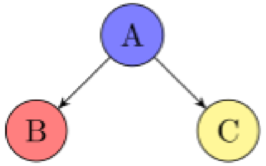
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

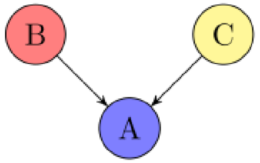
$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



$A \perp C \mid B$? Yes



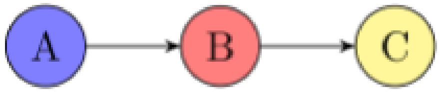
$B \perp C \mid A$? Yes



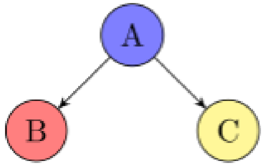
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

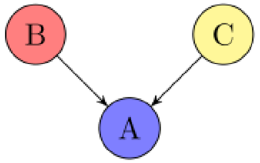
$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



$A \perp C \mid B$? Yes



$B \perp C \mid A$? Yes

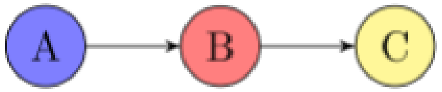


$B \perp C \mid A$?

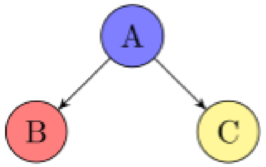
https://kunalmenda.com/2019/02/21/causation-and-correlation/

What does conditional independence mean?

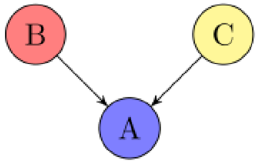
$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



$A \perp C \mid B$? Yes



$B \perp C \mid A$? Yes

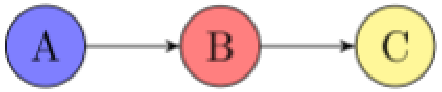


$B \perp C \mid A$? Inconclusive

https://kunalmenda.com/2019/02/21/causation-and-correlation/

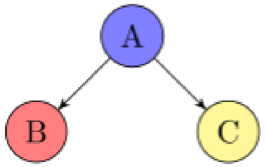
What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\qquad P(X \mid Z) = P(X \mid Y, Z)$

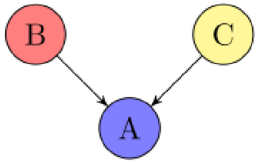


$A \perp C \mid B$? Yes

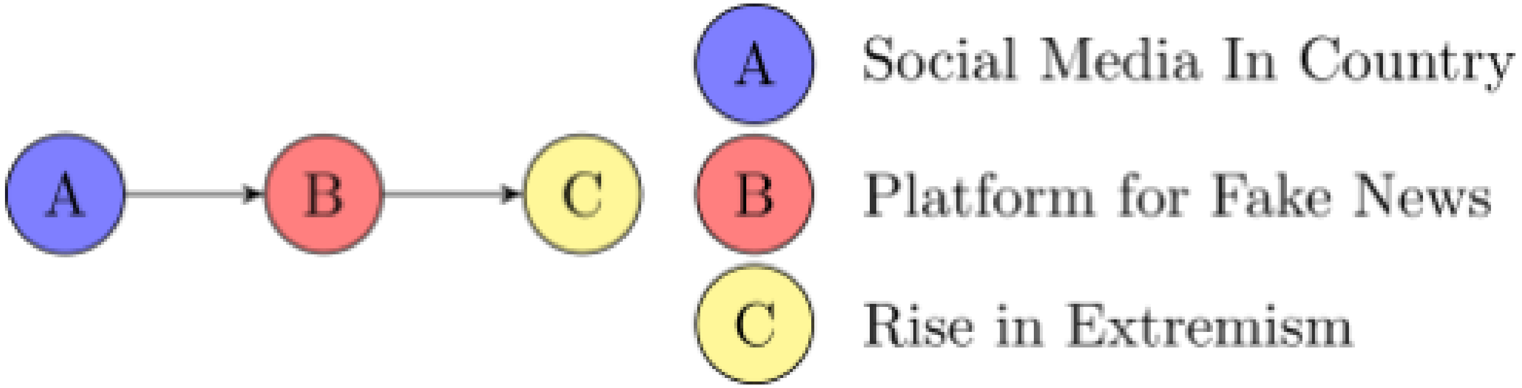
Mediator



$B \perp C \mid A$? Yes



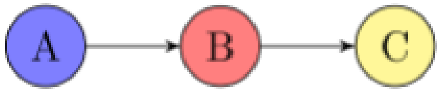
$B \perp C \mid A$? Inconclusive



https://kunalmenda.com/2019/02/21/causation-and-correlation/

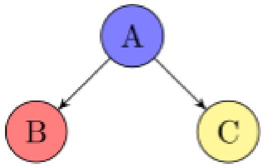
What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $\quad P(X \mid Z) = P(X \mid Y, Z)$



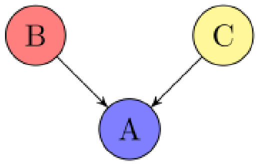
$A \perp C \mid B$? Yes

Mediator

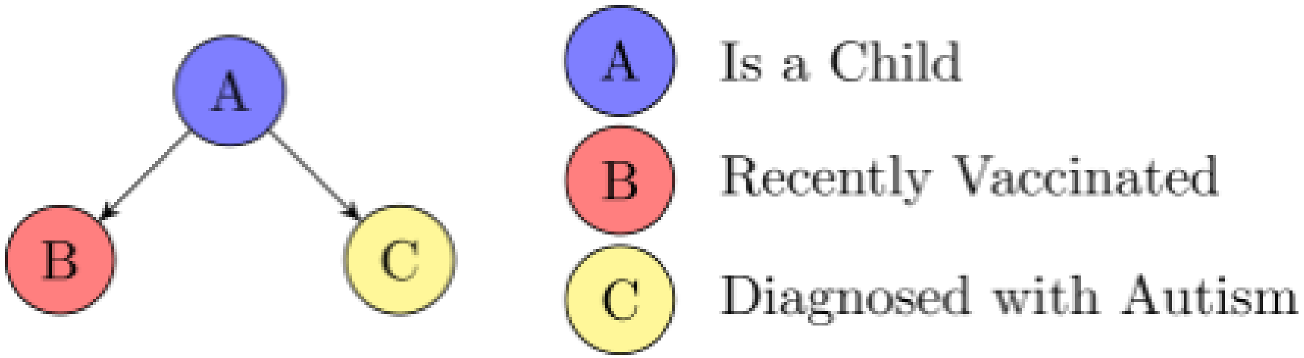


$B \perp C \mid A$? Yes

Confounder

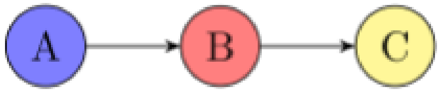


$B \perp C \mid A$? Inconclusive



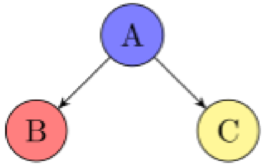
What does conditional independence mean?

$X \perp Y \mid Z \implies$ All of X's influence on Y comes through Z $P(X \mid Z) = P(X \mid Y, Z)$



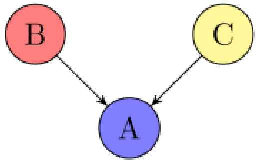
$A \perp C \mid B$? Yes

Mediator



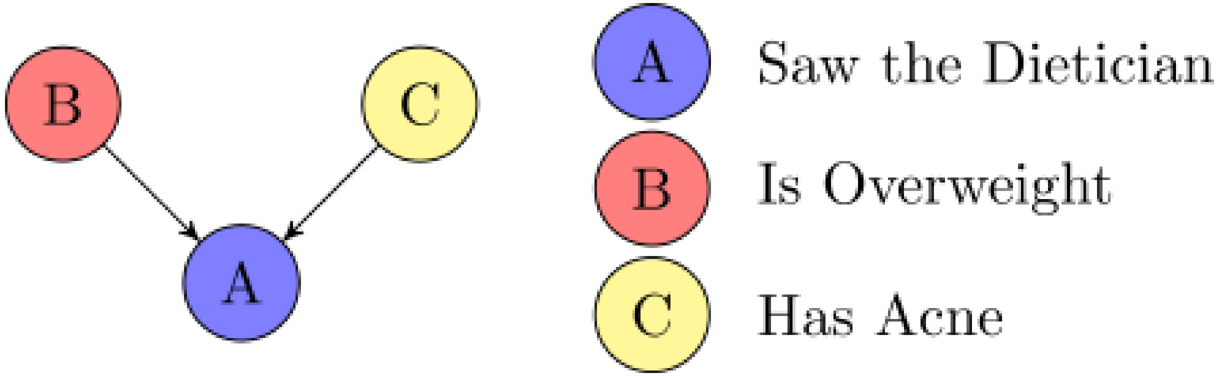
$B \perp C \mid A$? Yes

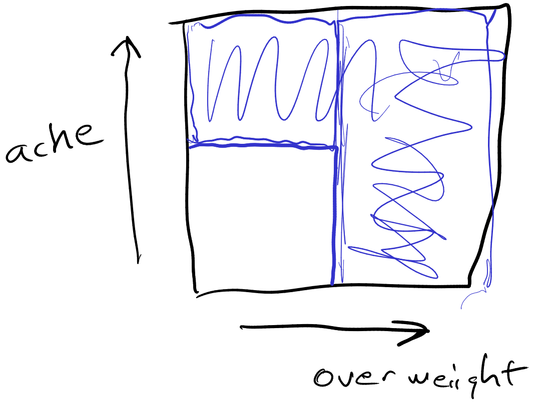
Confounder



$B \perp C \mid A$? Inconclusive

Collider

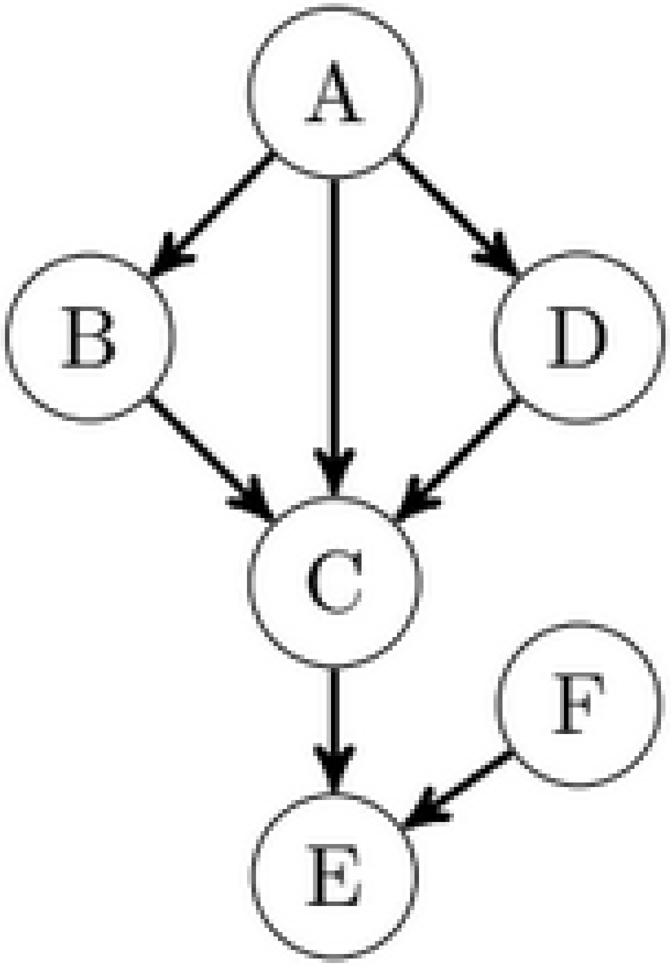




https://kunalmenda.com/2019/02/21/causation-and-correlation/

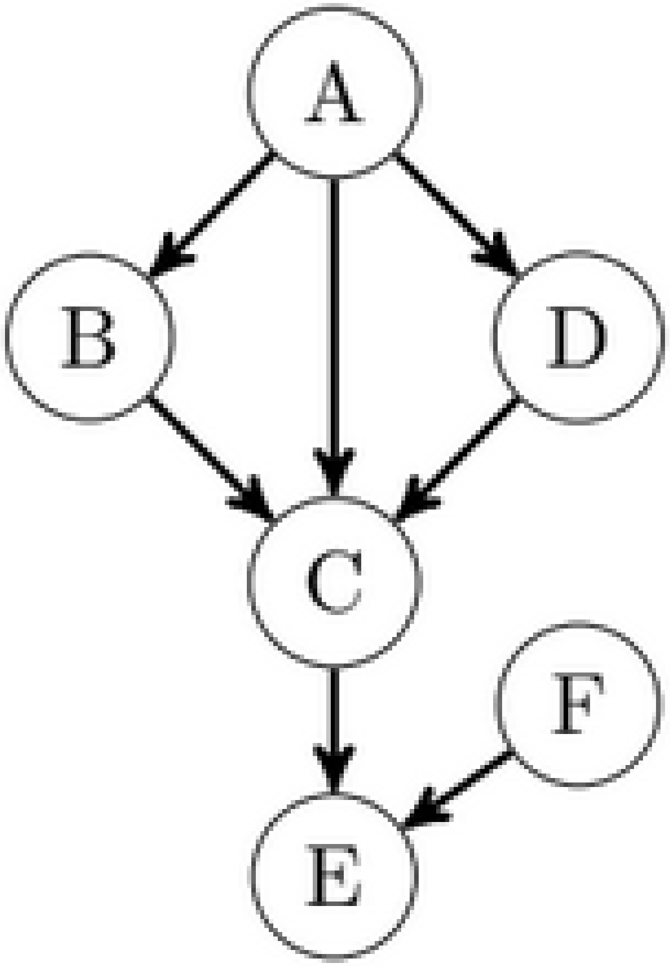
More Complex Example

More Complex Example

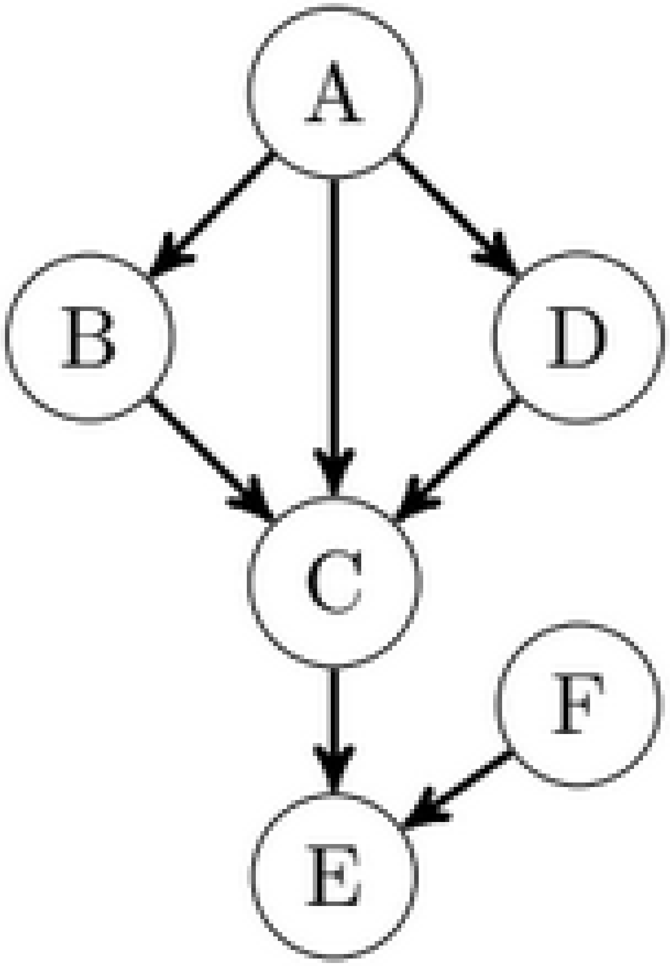


More Complex Example

$(B \perp D \mid A)$?



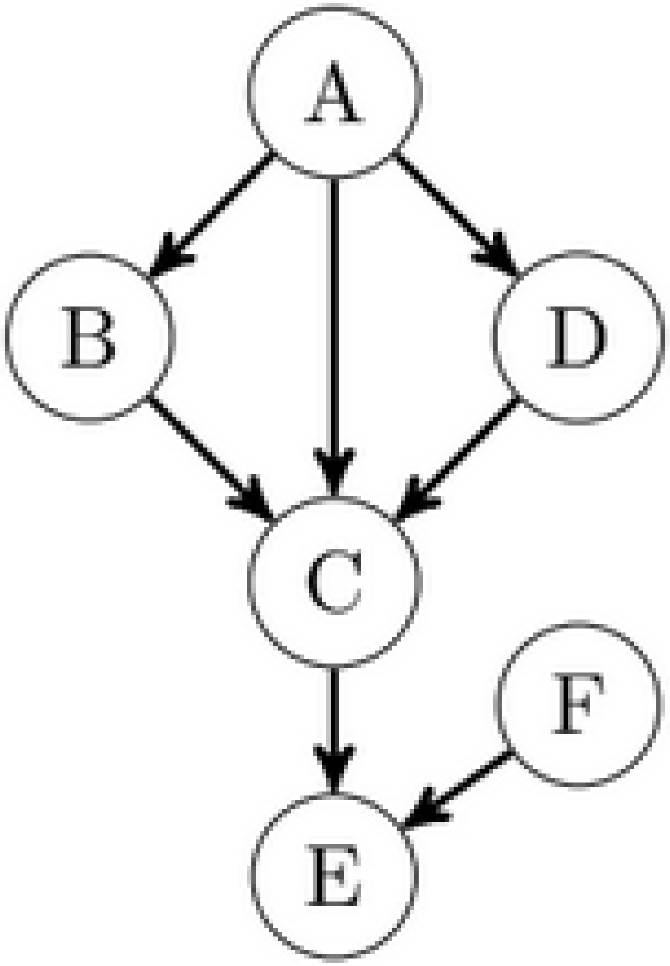
More Complex Example



$(B \perp D \mid A)$?

Yes!

More Complex Example

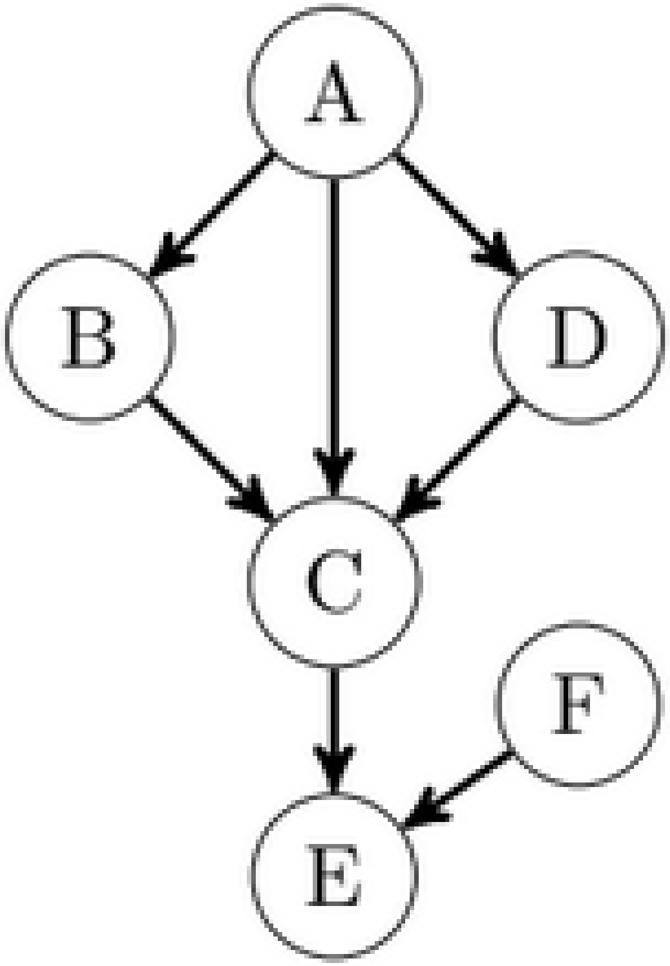


$(B \perp D \mid A)$?

Yes!

$(B \perp D \mid E)$?

More Complex Example



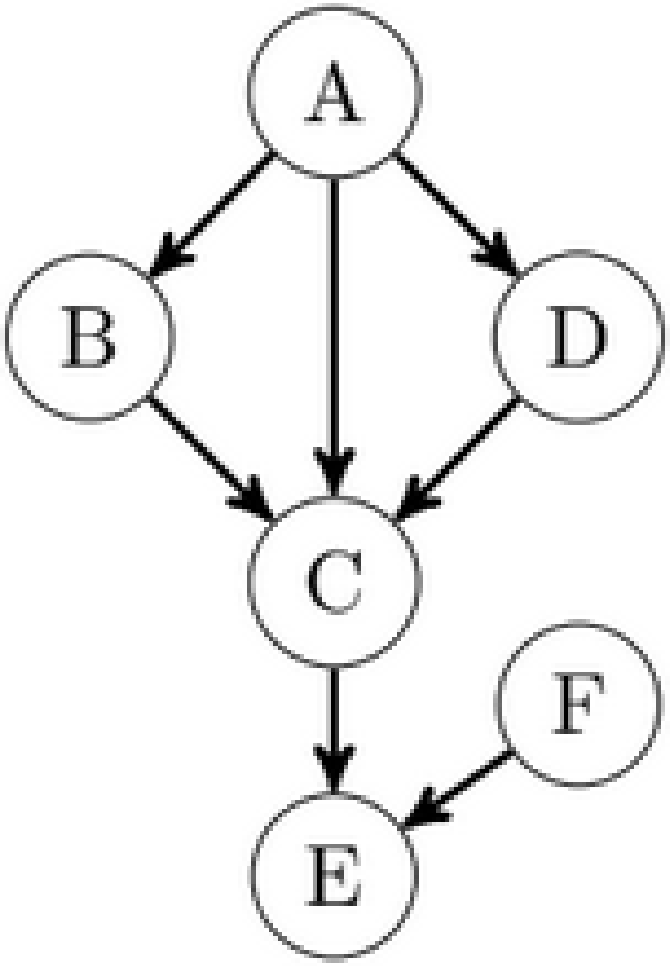
$(B \perp D \mid A)$?

Yes!

$(B \perp D \mid E)$?

Inconclusive

More Complex Example



$(B \perp D \mid A)$?

Yes!

$(B \perp D \mid E)$?

Inconclusive

Why is this relevant to decision making?

$A \perp B \mid C$

d-Separation

*short for "directionally separated"

d-Separation

A, B, C, D, E

$\mathcal{C} = \{C, D\}$

Let $\mathcal{C}$ be a set of random variables.

*short for "directionally separated"

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated** by $\mathcal{C}$ if any of the following are true

*short for "directionally separated"

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated** by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$

*short for "directionally separated"

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated** by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$

*short for "directionally separated"

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated** by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that Y is *not* in $\mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

*short for "directionally separated"

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated** by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that Y is *not* in $\mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

We say that A and B are *d-separated* by $\mathcal{C}$ if all paths between A and B are d-separated by $\mathcal{C}$.

*short for "directionally separated"

d-Separation

Let $\mathcal{C}$ be a set of random variables.

A *path* between A and B is *d-separated** by $\mathcal{C}$ if any of the following are true

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that Y is *not* in $\mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

We say that A and B are *d-separated* by $\mathcal{C}$ if all paths between A and B are d-separated by $\mathcal{C}$.

If A and B are d-separated by $\mathcal{C}$ then $A \perp B \mid \mathcal{C}$

*short for "directionally separated"

Proving Conditional Independence

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

Proving Conditional Independence

1. Enumerate all (non-cyclic) paths between nodes in question

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

Proving Conditional Independence

1. Enumerate all (non-cyclic) paths between nodes in question
2. Check all paths for d-separation

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

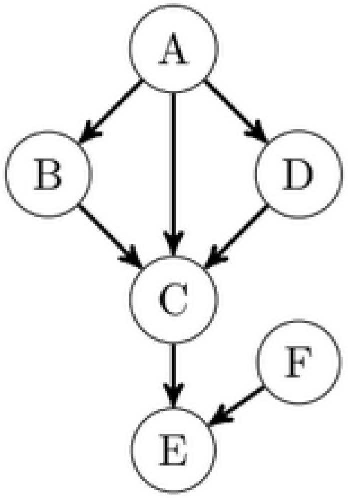
Proving Conditional Independence

1. Enumerate all (non-cyclic) paths between nodes in question
2. Check all paths for d-separation
3. If all paths d-separated, then CE

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

Proving Conditional Independence

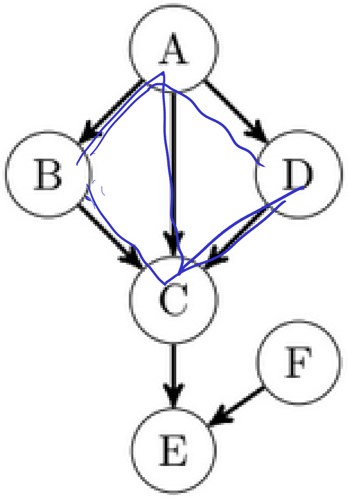
1. Enumerate all (non-cyclic) paths between nodes in question
2. Check all paths for d-separation
3. If all paths d-separated, then CE



1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

Proving Conditional Independence

✓ 1. Enumerate all (non-cyclic) paths between nodes in question
2. Check all paths for d-separation
3. If all paths d-separated, then CE ← conditionally independent



Example: $(B \perp D \mid \overbrace{C, E}^{\mathcal{C}})$?

None true → not d separated → inconclusive based only on structure

$B \leftarrow A \rightarrow D$

$B \rightarrow C \leftarrow D$ — rule 3 — not d-separated

$B \leftarrow A \rightarrow C \leftarrow D$ — rule 2 — not d-sep; rule 3 — not d-sep

$B \rightarrow C \leftarrow A \rightarrow D$ — rule 3 — not d-sep; rule 2 — not d-sep

No 1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$

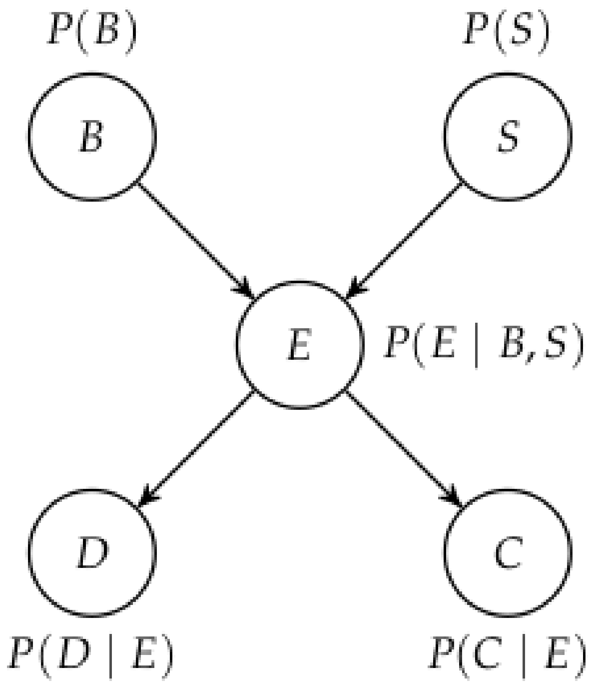
No 2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$

No 3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

Exercise

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

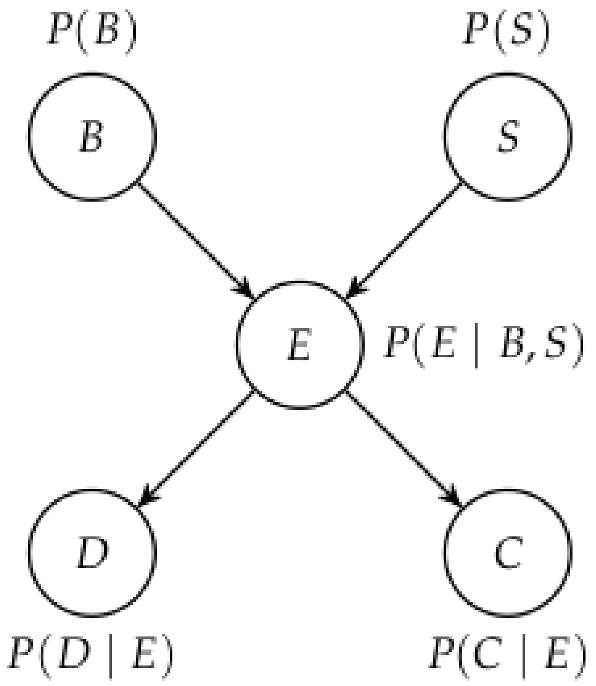
Exercise



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

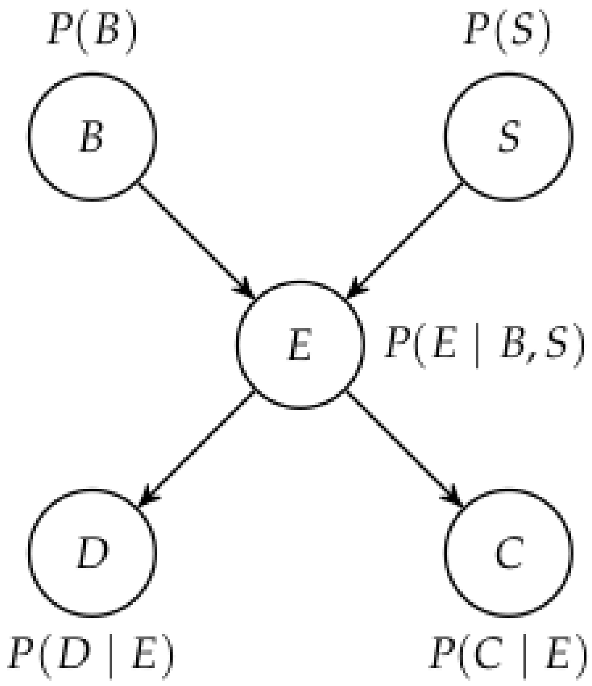
Exercise



B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

$$D \perp C \mid B \,?$$

1. The path contains a *chain* $X \rightarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \rightarrow Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \rightarrow Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.



Exercise

$$D \perp C \mid B \,?$$

$$D \perp C \mid E \,?$$

B battery failure
S solar panel failure
E electrical system failure
D trajectory deviation
C communication loss

1. The path contains a *chain* $X \to Y \to Z$ such that $Y \in \mathcal{C}$
2. The path contains a *fork* $X \leftarrow Y \to Z$ such that $Y \in \mathcal{C}$
3. The path contains an *inverted fork* (v-structure) $X \to Y \leftarrow Z$ such that $Y \notin \mathcal{C}$ and no descendant of Y is in $\mathcal{C}$.

Recap